TM- Exercise 2

2.1

(a) The function $f(\underline{x}) = \sqrt{x_1^2 + x_2^2}$ is not invertible since it is not injective: it's a projection from a 2D space to 1D space.

(b) $y = \sqrt{x_1^2 + x_2^2} + \varepsilon \Rightarrow \varepsilon = \sqrt{x_1^2 + x_2^2} - y$

$$\rightarrow p(y|x) = \frac{2}{\sqrt{2\pi}} \cdot \exp\left(- \frac{\left(\sqrt{x_1^2 + x_2^2} - y\right)^2}{2 \cdot \frac{1}{4}} \right)$$

(c) $\underline{y = 5}$, $p(x) = 1$

$$\rightarrow \underset{\underline{x}}{\operatorname{argmax}}\; p(x|y) = \underset{\underline{x}}{\operatorname{argmax}} \frac{p(y|x) \cdot p(x)}{p(y)} = \underset{\underline{x}}{\operatorname{argmax}}\; p(y|x)\, p(x)$$

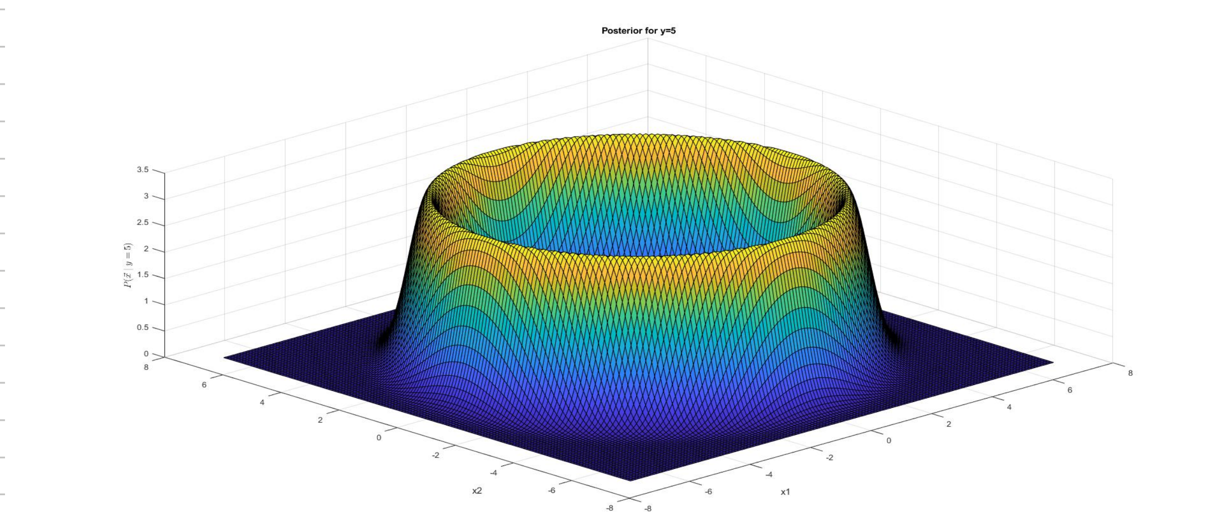
since does not depend on $p(y)$

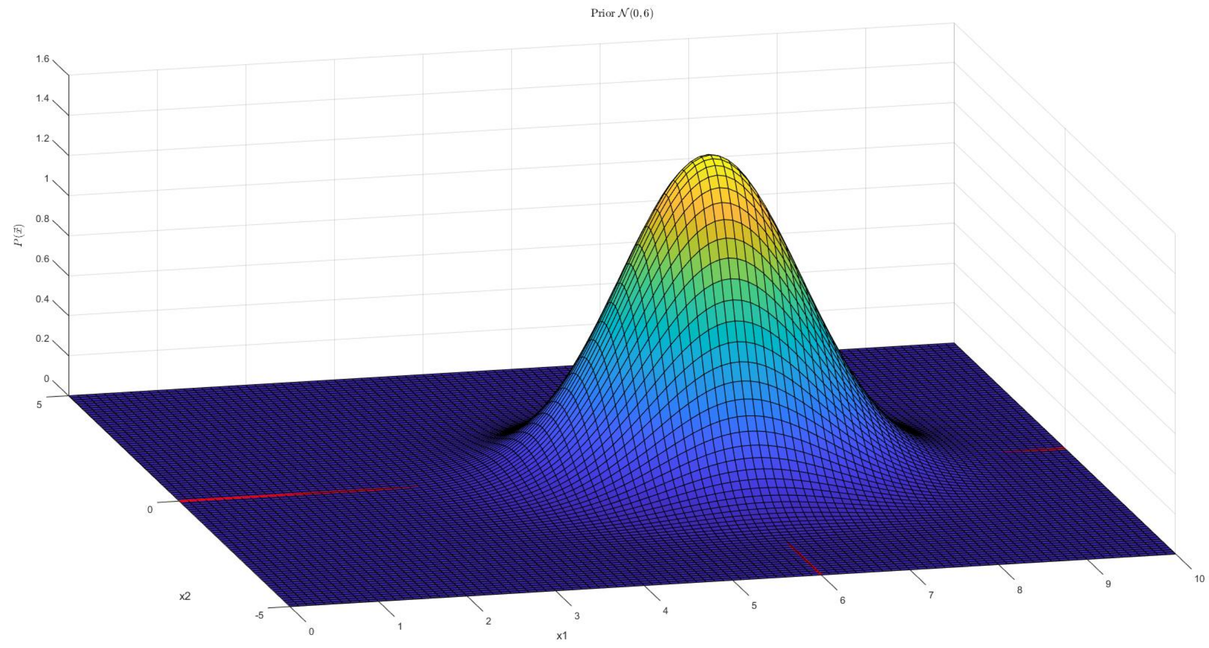
$$\rightarrow \underset{\underline{x}}{\operatorname{argmax}}\; p(y|x) \cdot 1 = \underset{x}{\operatorname{argmax}} \frac{2}{\sqrt{2\pi}} \cdot \exp\left(- \frac{\left(\sqrt{x_1^2 + x_2^2} - 5\right)^2}{1/2} \right)$$

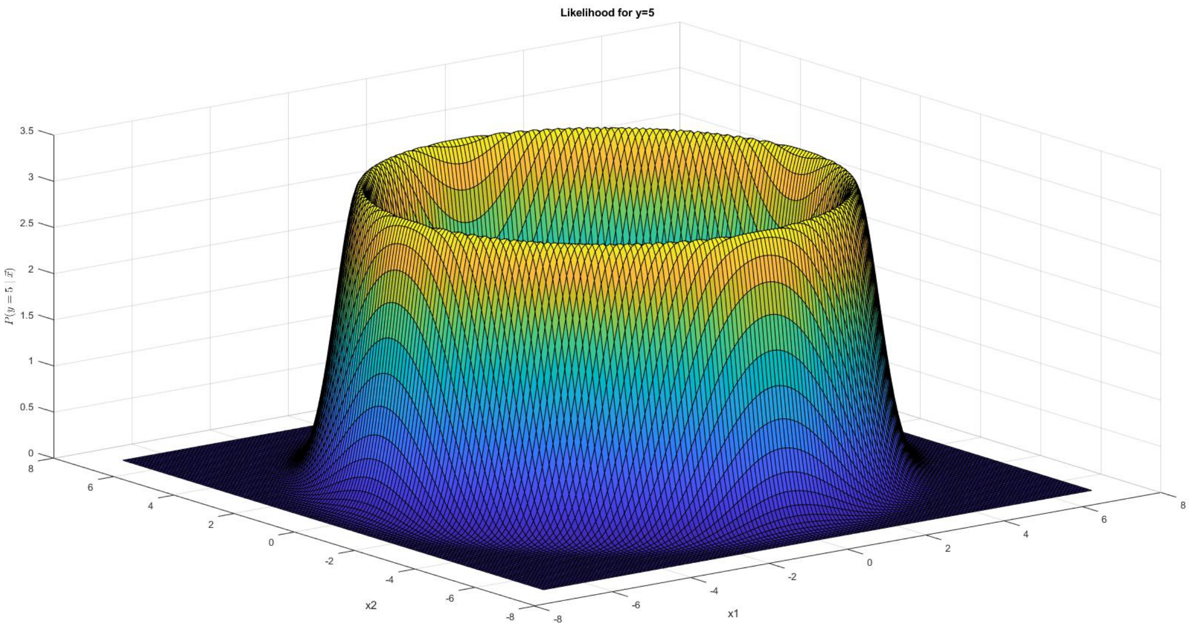
$$\rightarrow \max \text{ for } \sqrt{x_1^2 + x_2^2} = 5 \rightarrow \boxed{x_1^2 + x_2^2 = 25}$$

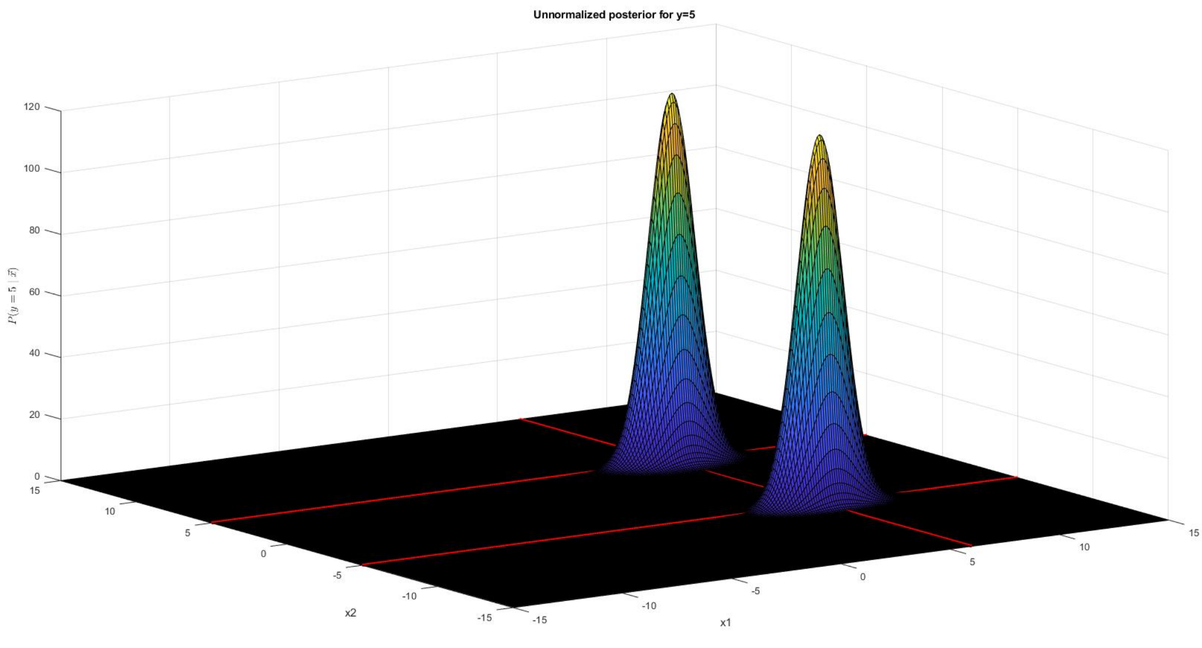
max value for all pairs of (x_1, x_2) sitting on the circle centered in 0 with radius 5.

(d)




Prior $\mathcal{N}(0, 6)$
$P(\vec{x})$
x1
x2


Likelihood for y=5
$P(y = 5 \mid \vec{x})$
x1
x2


Unnormalized posterior for y=5
$P(y = 5 \mid \vec{x})$
x1
x2

$\underline{\mu} = [6 \ \ 0]^T$

(f) $p(\underline{x} \mid y=5) \propto p(y=5 \mid \underline{x})\, p(\underline{x}) = \frac{2}{\sqrt{2\pi}} \exp\left(-2\left(\sqrt{x_1^2+x_2^2}-5\right)^2\right) \cdot$

$\cdot \frac{1}{2\pi} \cdot \exp\left(-\underbrace{[\underline{x}-\underline{\mu}]^T \Sigma^{-1} [x-\underline{\mu}]}_{\circledast}\right) \quad \circledast = (x_1-6)^2 + x_2^2$

$= \frac{2}{(2\pi)^{3/2}} \exp\left(-\left(2\left(\sqrt{x_1^2+x_2^2}-5\right)^2 + (x_1-6)^2 + x_2^2\right)\right)$

$\underset{\underline{x}}{\arg\max}\ p(\underline{x} \mid y=5) = \underset{\underline{x}}{\arg\min}\left[2\left(\sqrt{x_1^2+x_2^2}-5\right)^2 + (x_1-6)^2 + x_2^2\right] =$

$= \underset{x_2=0,\ x_1}{\arg\min}\left[\underbrace{2(x_1-5)^2 + (x_1-6)^2}_{(1)}\right]$

$\rightarrow \frac{d}{dx_1}(1) \overset{!}{=} 0 \rightarrow 2(2x_1-10) + 2x_1 - 12 = 0$

$\rightarrow 4x_1 - 20 + 2x_1 - 12 = 0 \rightarrow 6x_1 = 32 \rightarrow x_1 \approx 5{,}33$

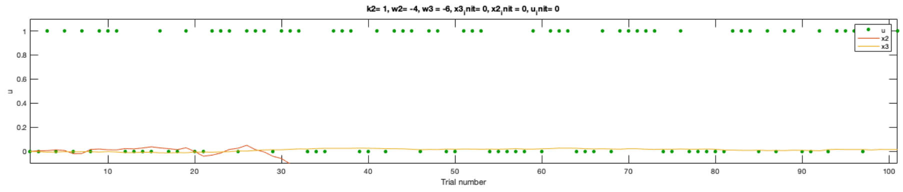
$\rightarrow \text{MAP} = (\hat{x}_1, \hat{x}_2) \approx (5.33, 0)$

(g)

Table of Contents

b) Generating the inputs with the generative model for the three level HGF

```
k2 = 1;
w2 = -4;
w3 = -6;
x3_init = 0;
x2_init = 0;
u_init = 0;
inputs = generate_inputs(k2,w2,w3,x3_init,x2_init,u_init);
u = inputs(:,1);
x2 = inputs(:,2);
x3 = inputs(:,3);
scrsz = get(0,'ScreenSize');
outerpos = [0.2*scrsz(3),0.7*scrsz(4),0.8*scrsz(3),0.3*scrsz(4)];
figure('OuterPosition', outerpos)
plot(u, '.', 'Color', [0 0.6 0], 'MarkerSize', 11)
xlabel('Trial number')
ylabel('u')
axis([1, length(inputs), -0.1, 1.1])
hold on;
plot(x2);
plot(x3);
legend('u','x2','x3')
str = sprintf('k2= %0.5g, w2= %0.5g, w3 = %0.5g, x3_init= %0.5g,
 x2_init = %0.5g, u_init= %0.5g', k2,w2,w3,x3_init,x2_init,u_init);
title(str)
hold off;
```

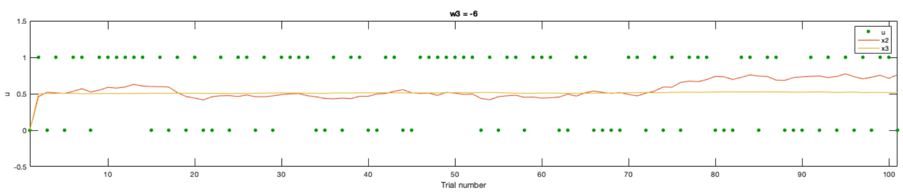


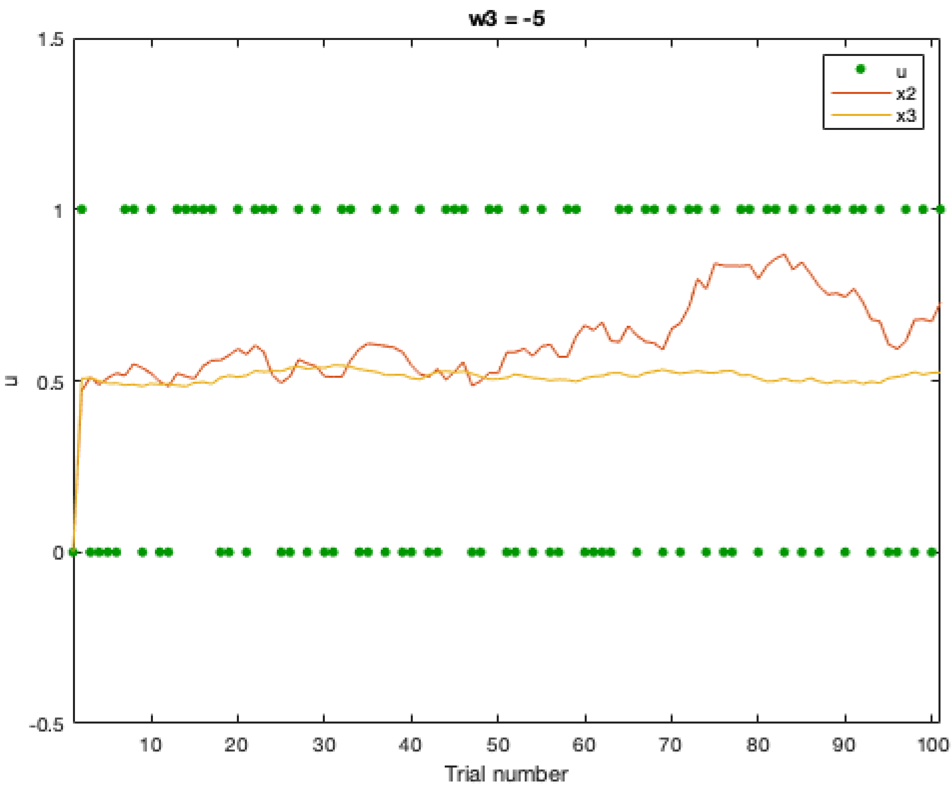
Trying out different thetas

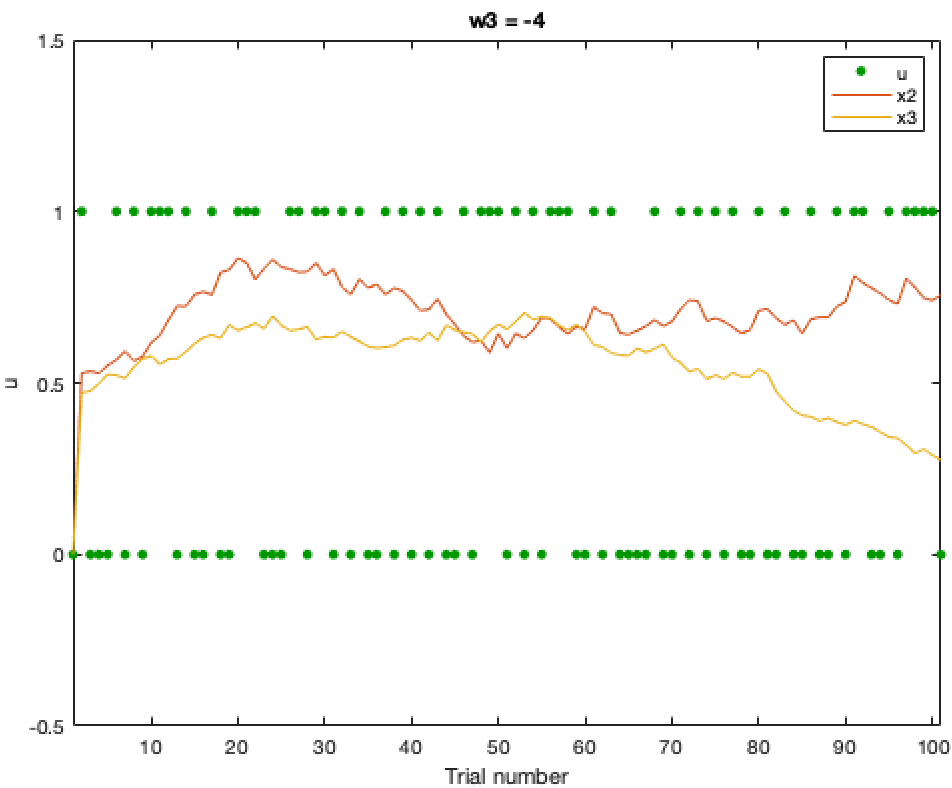
Higher volatility coefficients make the generated x2 and x3 much more variant. If it is too low, x2 (the tendency towards 1) becomes constant

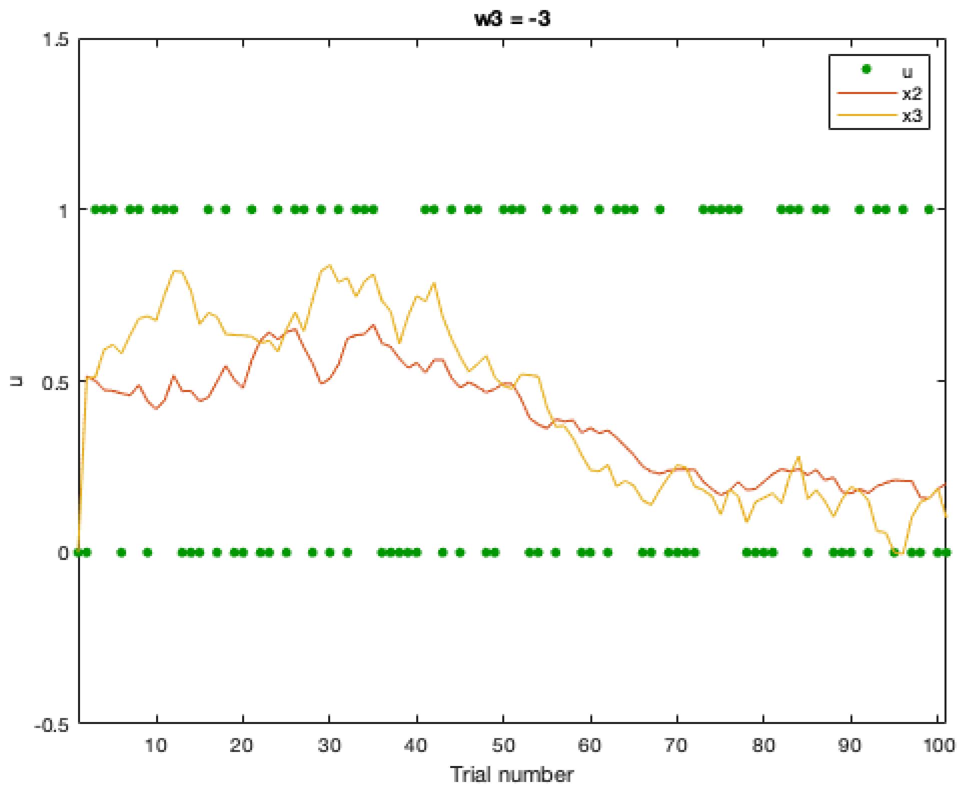
```
k2 = 1;
```

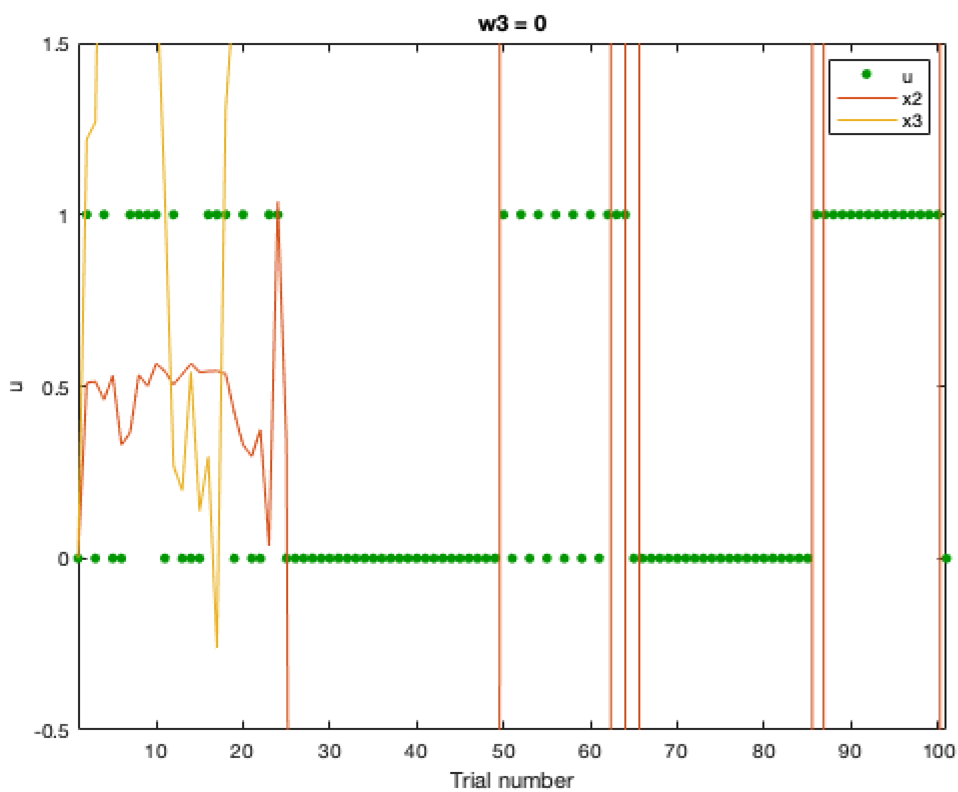
```
w2 = -4;
w3_list = [-6, -5, -4, -3, 0, 1];
x3_init = 0.5;
x2_init = 0.5;
u_init = 0;
it = 1;
for w3 = w3_list
inputs = generate_inputs(k2,w2,w3,x3_init,x2_init,u_init);
u = inputs(:,1);
x2 = inputs(:,2);
x3 = inputs(:,3);
figure(it)
plot(u, '.', 'Color', [0 0.6 0], 'MarkerSize', 11)
xlabel('Trial number')
ylabel('u')
axis([1, length(inputs), -0.5, 1.5])
hold on;
plot(x2);
plot(x3);
legend('u','x2','x3')
str = sprintf('w3 = %d', w3);
title(str)
hold off;
it = it + 1;
end
```

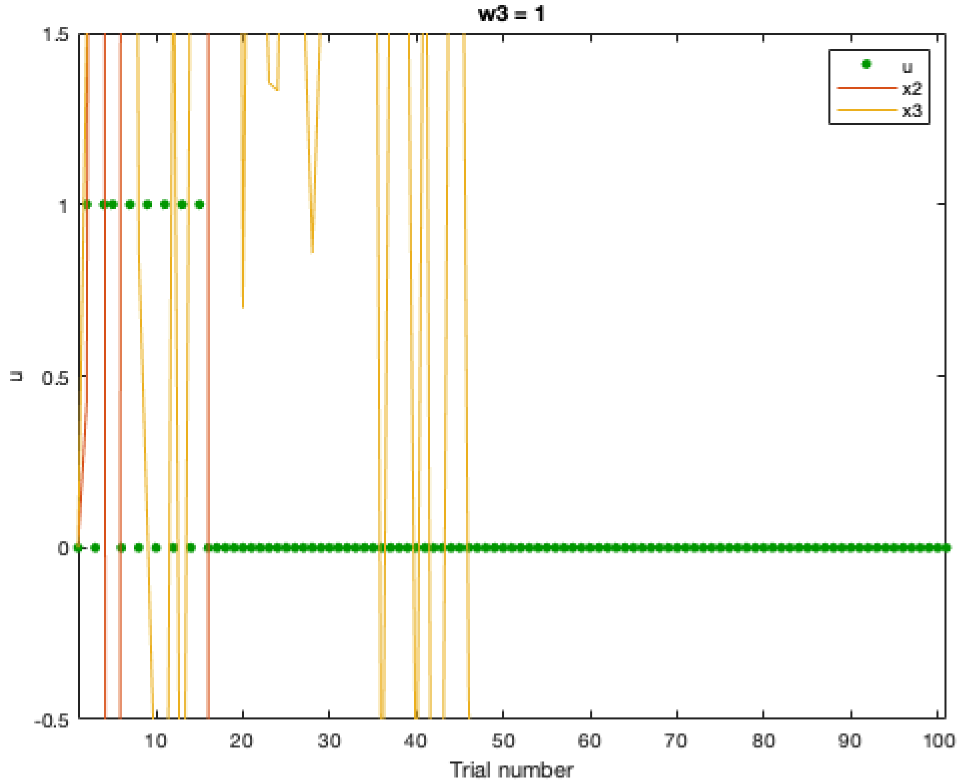



w3 = -5
u
x2
x3
1.5
1
0.5
0
-0.5
10
20
30
40
50
60
70
80
90
100
Trial number


w3 = -4
u
x2
x3
1.5
1
0.5
0
-0.5
10
20
30
40
50
60
70
80
90
100
Trial number


w3 = -3
1.5
1
0.5
0
-0.5
u
x2
x3
u
10
20
30
40
50
60
70
80
90
100
Trial number

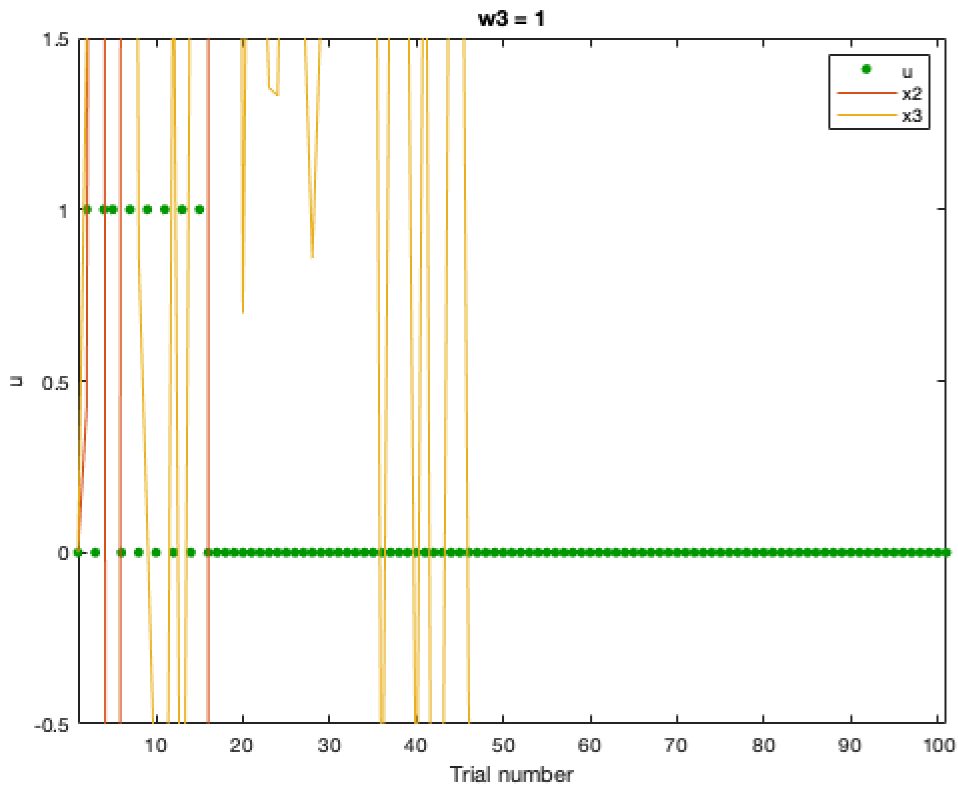

w3 = 0
1.5
1
0.5
0
-0.5
u
x2
x3
u
10
20
30
40
50
60
70
80
90
100
Trial number

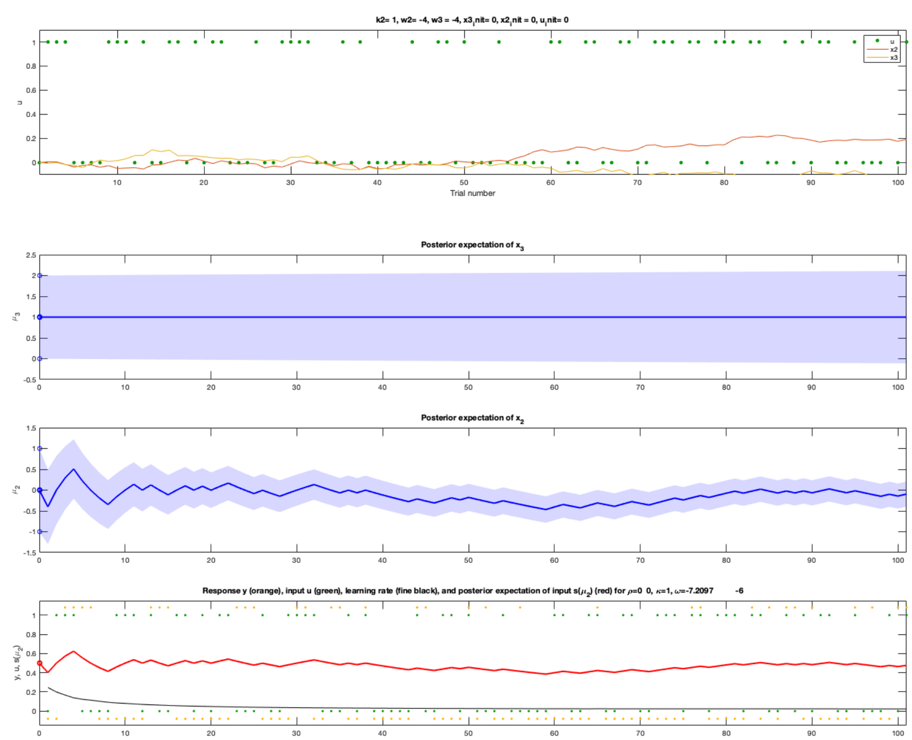


c) Simulating beliefs and responses

```
%The estimates for w2 and w3 (-7.2097, -6.0000) are far of the
 original values (-4,
%-4).
%The simulated agent does not track well x3 (which is estimated
 constant at 1 but is in fact variant around 0), nor x2. In our
 simulation x2 varies between 0 and 0.2 while the agent estimates the
 x2 between -0.5 and 0.5.

addpath('../tapas/HGF')
k2 = 1;
w2 = -4;
w3 = -4;
x3_init = 0;
x2_init = 0;
u_init = 0;
inputs = generate_inputs(k2,w2,w3,x3_init,x2_init,u_init);
u = inputs(:,1);
x2 = inputs(:,2);
x3 = inputs(:,3);

scrsz = get(0,'ScreenSize');
outerpos = [0.2*scrsz(3),0.7*scrsz(4),0.8*scrsz(3),0.3*scrsz(4)];
figure('OuterPosition', outerpos)
```

```
plot(u, '.', 'Color', [0 0.6 0], 'MarkerSize', 11)
xlabel('Trial number')
ylabel('u')
axis([1, length(inputs), -0.1, 1.1])
hold on;
plot(x2);
plot(x3);
legend('u','x2','x3')
str = sprintf('k2= %0.5g, w2= %0.5g, w3 = %0.5g, x3_init= %0.5g,
 x2_init = %0.5g, u_init= %0.5g', k2,w2,w3,x3_init,x2_init,u_init);
title(str)
hold off;

est = tapas_fitModel([],...
                         u,...
                         'tapas_hgf_binary_config',...
                         'tapas_bayes_optimal_binary_config',...
                         'tapas_quasinewton_optim_config');
sim = tapas_simModel(u,...
                     'tapas_hgf_binary',...
                     [NaN 0 1 NaN 1 1 NaN 0 0 1 1 NaN
 est.optim.final(13) est.optim.final(14)],...
                     'tapas_unitsq_sgm',...
                     5,...
                     12345);

tapas_hgf_binary_plotTraj(sim)

Ignored trials: none
Irregular trials: none

Optimizing...

Calculating the log-model evidence (LME)...

Results:

Parameter estimates for the perceptual model:
    mu_0: [NaN 0 1]
    sa_0: [NaN 0.1000 1]
     rho: [NaN 0 0]
      ka: [1 1]
      om: [NaN -7.2097 -6.0000]

Model quality:
    LME (more is better): -72.0156
    AIC (less is better): 145.609
    BIC (less is better): 150.8393

    AIC and BIC are approximations to -2*LME = 144.0313.

Ignored trials: none
```


w3 = 1
u
x2
x3
Trial number


k2= 1, w2= -4, w3 = -4, x3_init= 0, x2_init = 0, u_init= 0
u
x2
x3
Trial number
Posterior expectation of x_3
μ_3
Posterior expectation of x_2
μ_2
Response y (orange), input u (green), learning rate (fine black), and posterior expectation of input s(μ_2) (red) for ρ=0 0, κ=1, ω=-7.2097 -6
y, u, s(μ_2)
Trial number

```
close all
```

Published with MATLAB® R2018b

Task 2.3

Table of Contents


```
addpath('tapas/HGF')
clear all;
close all;

u = load('example_binary_input.txt');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

Warning: Name is nonexistent or not a directory: /Users/Hendrik/Documents/MATLAB/TNU/Exercise2/tapas/HGF

a)k2=2.5, w2=-4, w3=-6, mu3=1, sa3=1, ze =5

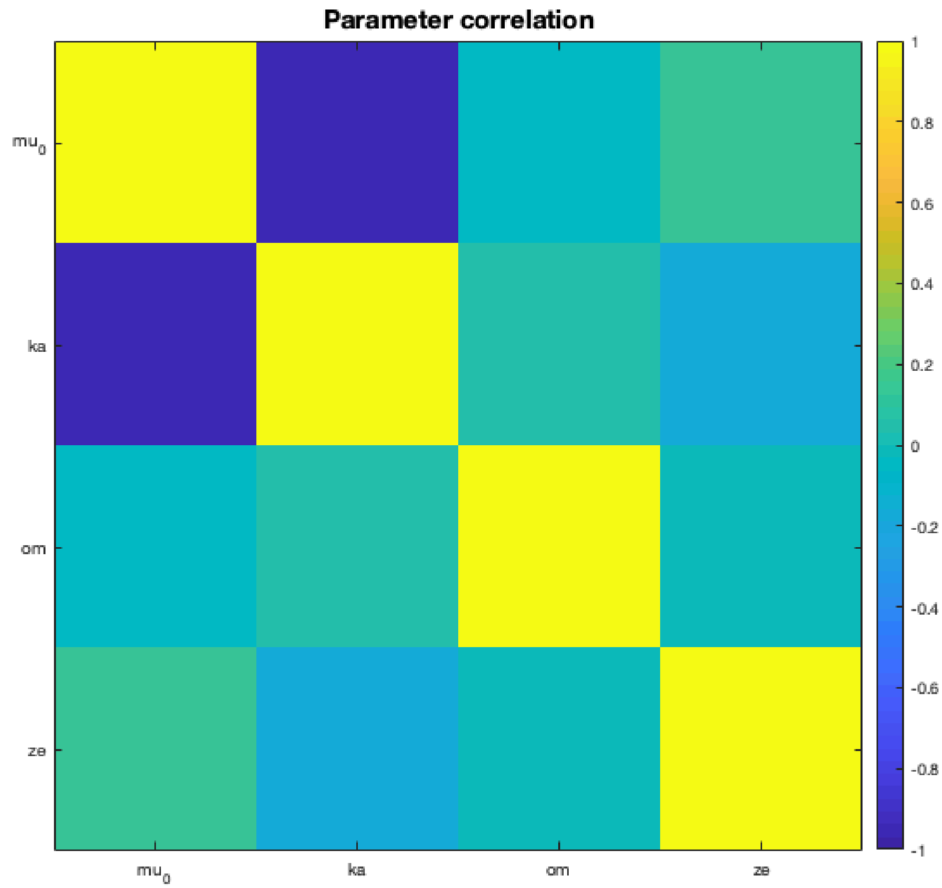
```
% The \omega_3 is very accurate in all our results.
% The estimated parameters are not as accurate as the set ones since
 the
% variance is not zero. For example the \kappa_2 is not accurate when
 we
% estimate it but nearly 2.5 when not estimating this parameter.

% The first covariance plot shows large correlation between \kappa_2
 and
% \mu_0.
% The second covariance plot shows large correlation between \omega_2
 and
% \mu_0.


%simulate model
sim = tapas_simModel(u,...
'tapas_hgf_binary',...
[NaN 0 1 NaN 1 1 NaN 0 0 1 2.5 NaN -4 -6],...
'tapas_unitsq_sgm',...
5);

%estimate param: (ze,mu3,K2,exp(w3))
est1 = tapas_fitModel(sim.y,...
```

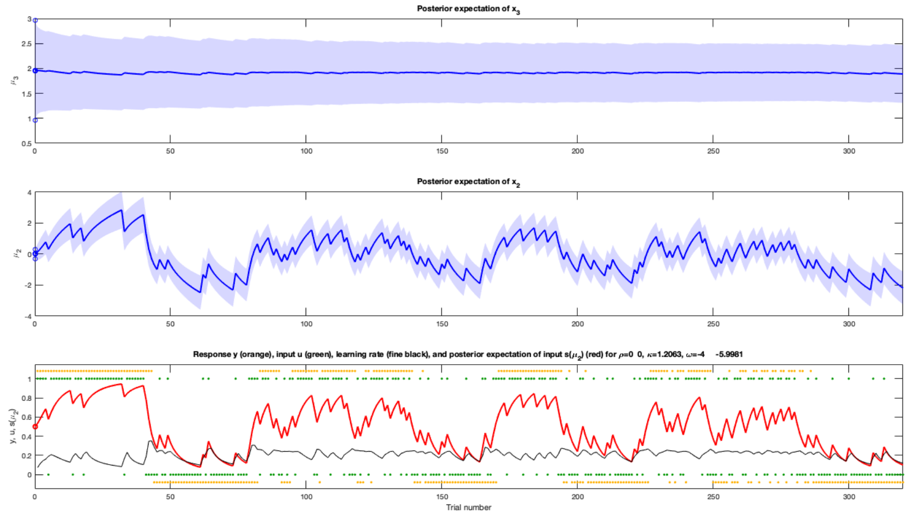
```
                     sim.u,...
                     'tapas_hgf_binary_config_2',...
                     'tapas_unitsq_sgm_config',...
                     'tapas_quasinewton_optim_config')


%plot posterior correlation
tapas_fit_plotCorr(est1)

%plot trajectories
tapas_hgf_binary_plotTraj(est1)


%estimate param: (ze,mu3,w2,exp(w3))
est2 = tapas_fitModel(sim.y,...
                     sim.u,...
                     'tapas_hgf_binary_config_3',...
                     'tapas_unitsq_sgm_config',...
                     'tapas_quasinewton_optim_config')


%plot posterior correlation
tapas_fit_plotCorr(est2)

%plot trajectories
tapas_hgf_binary_plotTraj(est2)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%
```

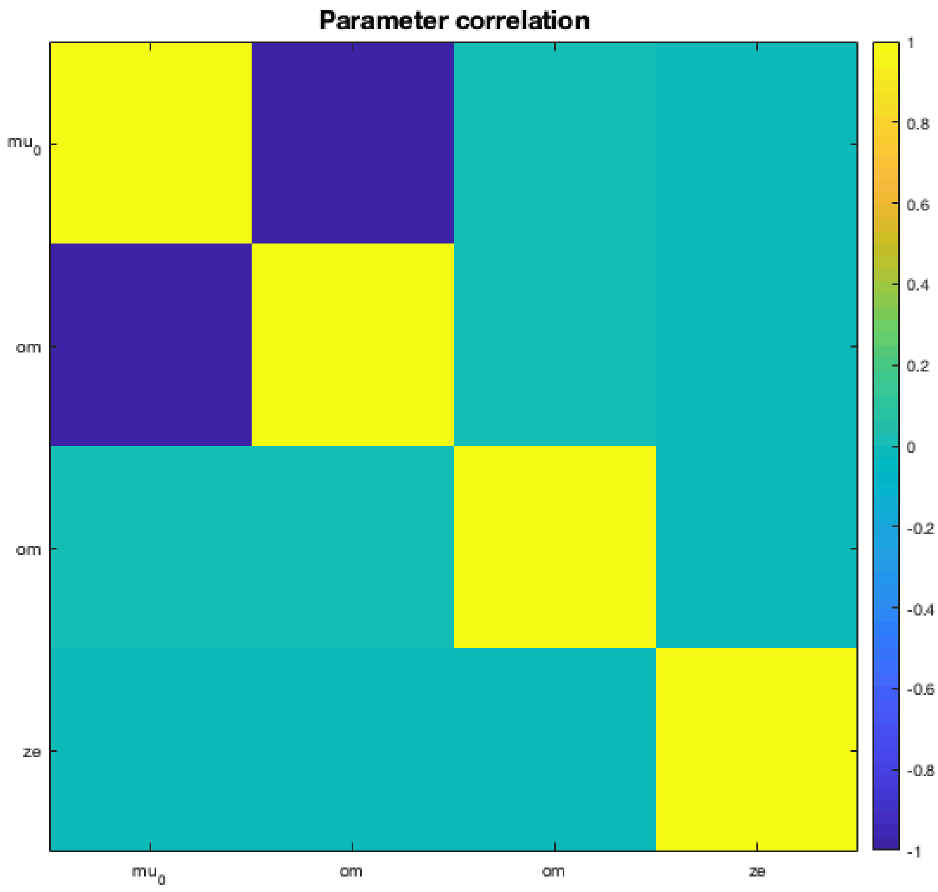
```
Ignored trials: none
Ignored trials: none
Irregular trials: none

Optimizing...

Calculating the log-model evidence (LME)...

Results:

Parameter estimates for the perceptual model:
    mu_0: [NaN 0 1.9604]
    sa_0: [NaN 0.1000 1]
     rho: [NaN 0 0]
      ka: [1 1.2063]
      om: [NaN -4 -5.9981]


Parameter estimates for the observation model:
    ze: 7.1749

Model quality:
    LME (more is better): -48.0735
    AIC (less is better): 90.2843
```

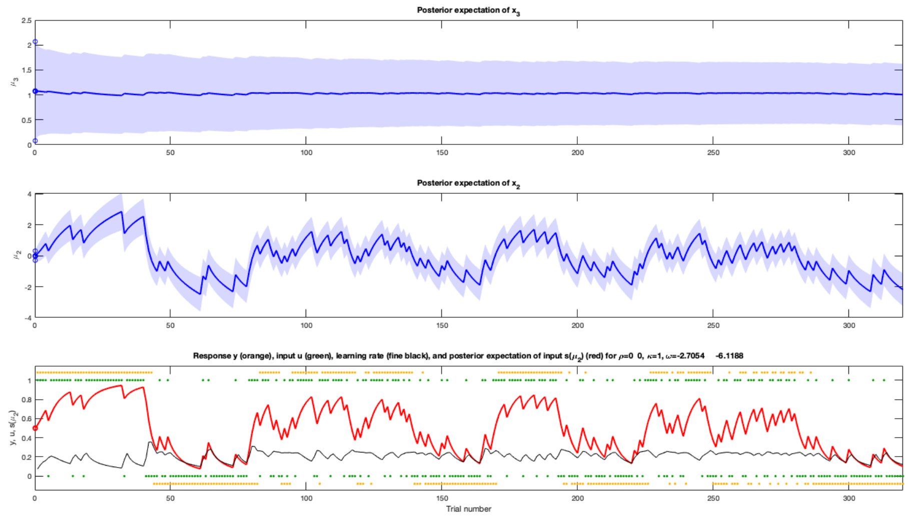
```
    BIC (less is better): 105.3576

    AIC and BIC are approximations to -2*LME = 96.1471.


est1 = 

  struct with fields:

        y: [320×1 double]
        u: [320×1 double]
      ign: []
      irr: [0×1 double]
    c_prc: [1×1 struct]
    c_obs: [1×1 struct]
    c_opt: [1×1 struct]
    optim: [1×1 struct]
    p_prc: [1×1 struct]
    p_obs: [1×1 struct]
     traj: [1×1 struct]

Ignored trials: none
Irregular trials: none

Optimizing...

Calculating the log-model evidence (LME)...

Results:

Parameter estimates for the perceptual model:
    mu_0: [NaN 0 1.0747]
    sa_0: [NaN 0.1000 1]
     rho: [NaN 0 0]
      ka: [1 1]
      om: [NaN -2.7054 -6.1188]


Parameter estimates for the observation model:
    ze: 7.1569

Model quality:
    LME (more is better): -48.1332
    AIC (less is better): 90.281
    BIC (less is better): 105.3543

    AIC and BIC are approximations to -2*LME = 96.2663.


est2 = 

  struct with fields:

        y: [320×1 double]
```

```
    u: [320×1 double]
  ign: []
  irr: [0×1 double]
c_prc: [1×1 struct]
c_obs: [1×1 struct]
c_opt: [1×1 struct]
optim: [1×1 struct]
p_prc: [1×1 struct]
p_obs: [1×1 struct]
 traj: [1×1 struct]
```






Parameter correlation
mu_0
om
om
ze
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1



b)k2=1, w2=-4, w3=-4.1674, mu3=2.5, sa3=6.25, ze = 5

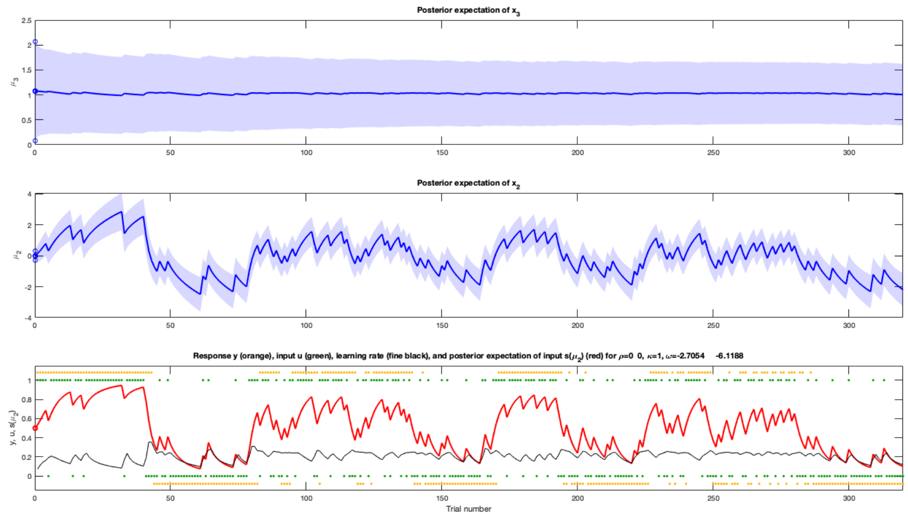
```
% The \omega_3 is very accurate in all our results.
% The estimated parameters are not as accurate as the set ones since
 the
% variance is not zero. For example the \kappa_2 is not accurate when
 we
% estimate it but equal to 1 when not estimating it.

% The estimates for \zeta and \theta are more accurate in b) than in
 a).
% We think a possible explanation for the better estimates is that the
% volatility coeffitient \theta is higher and therefore the model is
 more
% flexible

% The first covariance plot shows large correlation between \kappa_2
 and
% \mu_0.
% The second covariance plot shows large correlation between \omega_2
 and
% \mu_0.

%simulate model
sim2 = tapas_simModel(u,...
'tapas_hgf_binary',...
[NaN 0 2.5 NaN 1 6.25 NaN 0 0 1 1 NaN -4 -4.1674],...
'tapas_unitsq_sgm',...
5);

%estimate model
est3 = tapas_fitModel(sim2.y,...
```

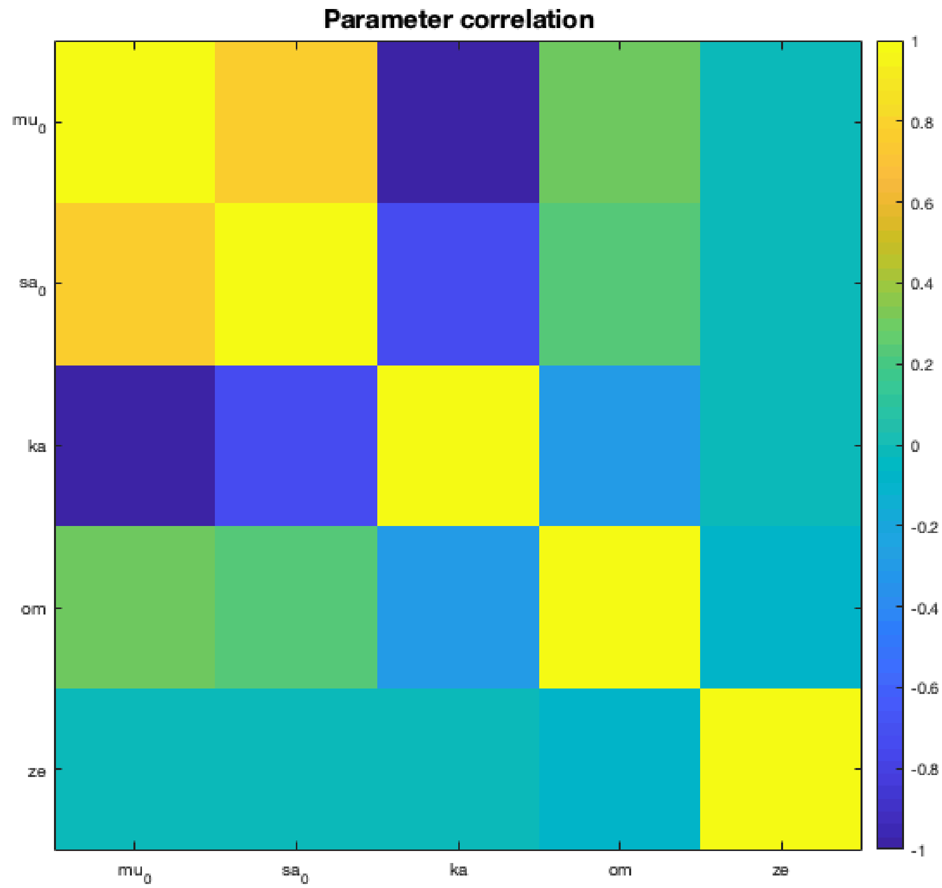
```
                     sim2.u,...
                     'tapas_hgf_binary_config_4',...
                     'tapas_unitsq_sgm_config',...
                     'tapas_quasinewton_optim_config')


%plot posterior correlation
tapas_fit_plotCorr(est3)

%plot trajectories
tapas_hgf_binary_plotTraj(est3)

%estimate model
est4 = tapas_fitModel(sim2.y,...
                     sim2.u,...
                     'tapas_hgf_binary_config_5',...
                     'tapas_unitsq_sgm_config',...
                     'tapas_quasinewton_optim_config')


%plot posterior correlation
tapas_fit_plotCorr(est4)

%plot trajectories
tapas_hgf_binary_plotTraj(est4)
```

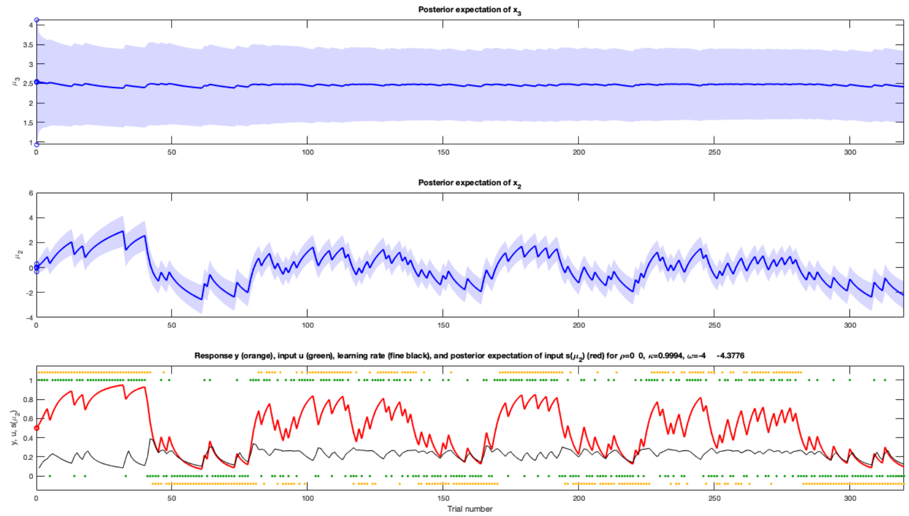
```
Ignored trials: none
Ignored trials: none
Irregular trials: none

Optimizing...

Calculating the log-model evidence (LME)...

Results:

Parameter estimates for the perceptual model:
    mu_0: [NaN 0 2.5364]
    sa_0: [NaN 0.1000 2.5626]
     rho: [NaN 0 0]
      ka: [1 0.9994]
      om: [NaN -4 -4.3776]


Parameter estimates for the observation model:
    ze: 4.7421

Model quality:
    LME (more is better): -67.6089
    AIC (less is better): 130.7859
    BIC (less is better): 149.6275

    AIC and BIC are approximations to -2*LME = 135.2177.
```

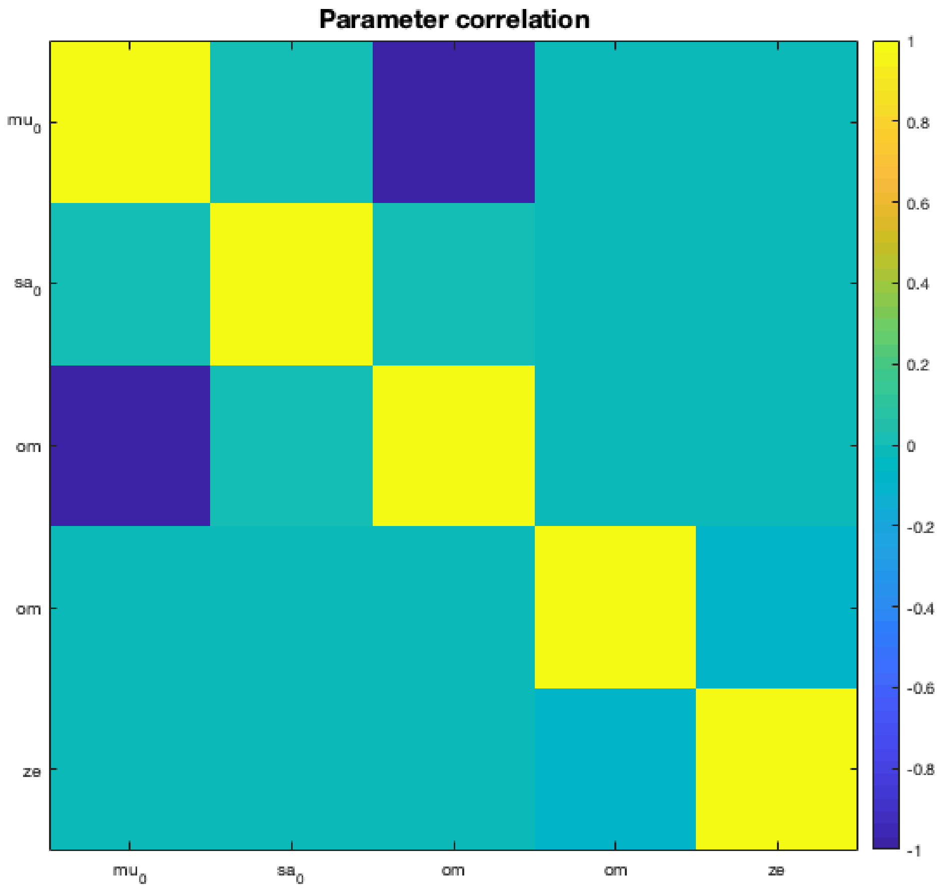
```
est3 =

  struct with fields:

        y: [320×1 double]
        u: [320×1 double]
      ign: []
      irr: [0×1 double]
    c_prc: [1×1 struct]
    c_obs: [1×1 struct]
    c_opt: [1×1 struct]
    optim: [1×1 struct]
    p_prc: [1×1 struct]
    p_obs: [1×1 struct]
     traj: [1×1 struct]

Ignored trials: none
Irregular trials: none

Optimizing...

Calculating the log-model evidence (LME)...

Results:

Parameter estimates for the perceptual model:
    mu_0: [NaN 0 2.5085]
    sa_0: [NaN 0.1000 2.5638]
     rho: [NaN 0 0]
      ka: [1 1]
      om: [NaN -3.9726 -4.5274]


Parameter estimates for the observation model:
    ze: 4.7457

Model quality:
    LME (more is better): -68.565
    AIC (less is better): 130.7821
    BIC (less is better): 149.6237

    AIC and BIC are approximations to -2*LME = 137.1299.


est4 =

  struct with fields:

        y: [320×1 double]
        u: [320×1 double]
      ign: []
      irr: [0×1 double]
    c_prc: [1×1 struct]
```

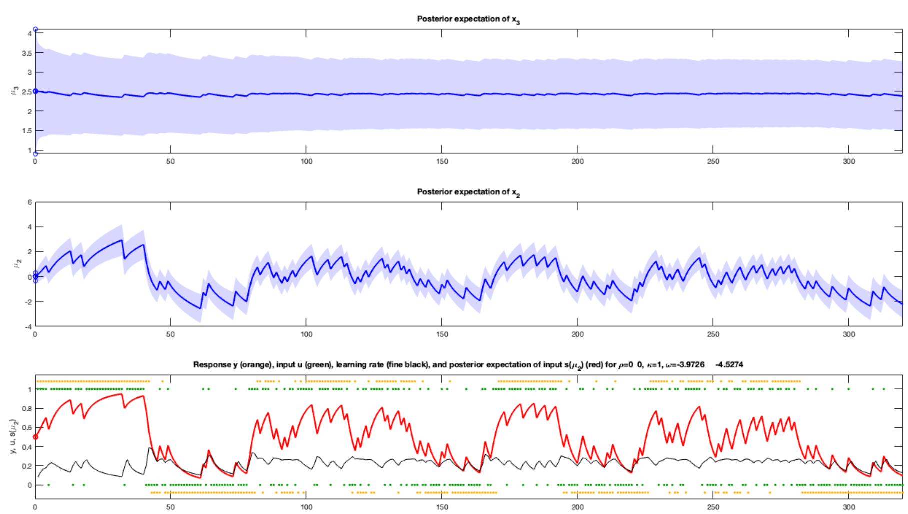
```
c_obs: [1×1 struct]
c_opt: [1×1 struct]
optim: [1×1 struct]
p_prc: [1×1 struct]
p_obs: [1×1 struct]
 traj: [1×1 struct]
```




Parameter correlation
mu_0
sa_0
ka
om
ze
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1


Posterior expectation of x_3
Posterior expectation of x_2
Response y (orange), input u (green), learning rate (fine black), and posterior expectation of input s(μ_2) (red) for ρ=0 0, κ=0.9994, ω=-4 -4.3776
y, u, s(μ_2)
Trial number


Parameter correlation
mu_0
sa_0
om
om
ze
mu_0
sa_0
om
om
ze
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1


Posterior expectation of x_3
Posterior expectation of x_2
Response y (orange), input u (green), learning rate (fine black), and posterior expectation of input s(μ_2) (red) for ρ=0 0, κ=1, ω=-3.9726 -4.5274
Trial number

c)

```
%The \mu_3 can only be changed without changing the other beliefs by
%changing \sigma_3 or \kappa_2.
```

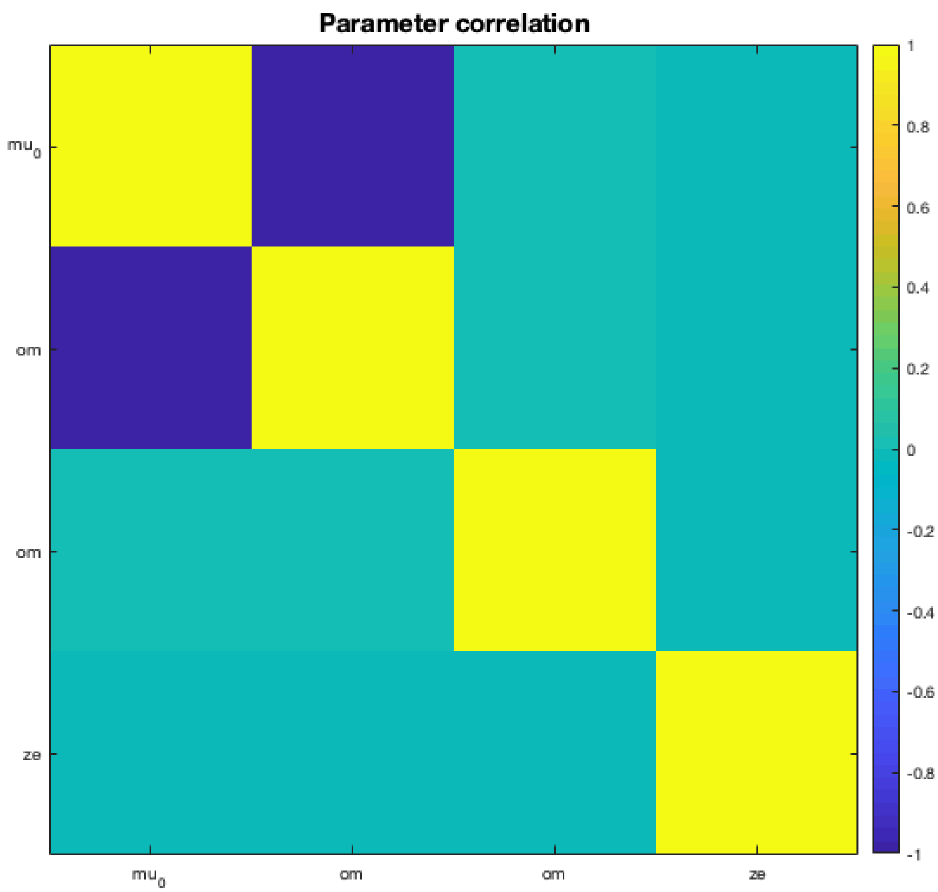
d) tapas_unitsq_sgm_mu3 as response model

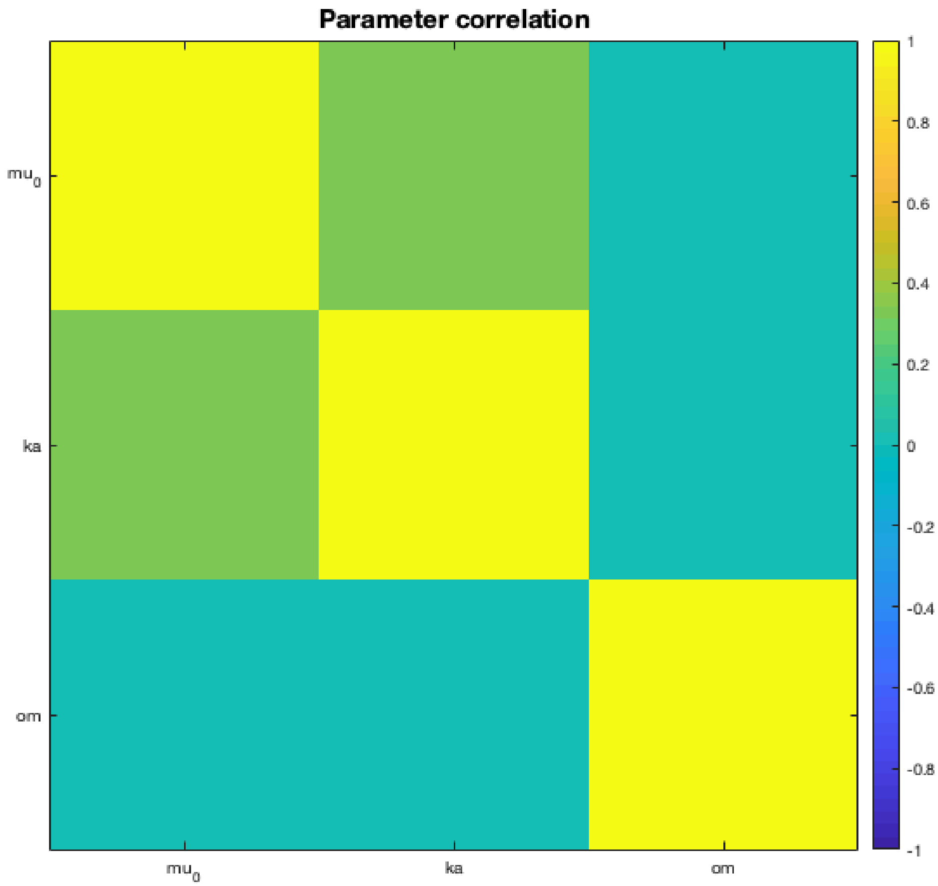
```
%simulate model
sim = tapas_simModel(u,...
'tapas_hgf_binary',...
[NaN 0 1 NaN 1 1 NaN 0 0 1 2.5 NaN -4 -6],...
'tapas_unitsq_sgm',...
5);

%estimate param: (ze,mu3,K2,exp(w3))
est1 = tapas_fitModel(sim.y,...
                     sim.u,...
                     'tapas_hgf_binary_config_2',...
                     'tapas_unitsq_sgm_mu3_config',...
                     'tapas_quasinewton_optim_config')


%plot posterior correlation
tapas_fit_plotCorr(est1)

%plot trajectories
tapas_hgf_binary_plotTraj(est1)


%estimate param: (ze,mu3,w2,exp(w3))
est2 = tapas_fitModel(sim.y,...
                     sim.u,...
                     'tapas_hgf_binary_config_3',...
                     'tapas_unitsq_sgm_mu3_config',...
                     'tapas_quasinewton_optim_config')


%plot posterior correlation
tapas_fit_plotCorr(est2)

%plot trajectories
tapas_hgf_binary_plotTraj(est2)
```

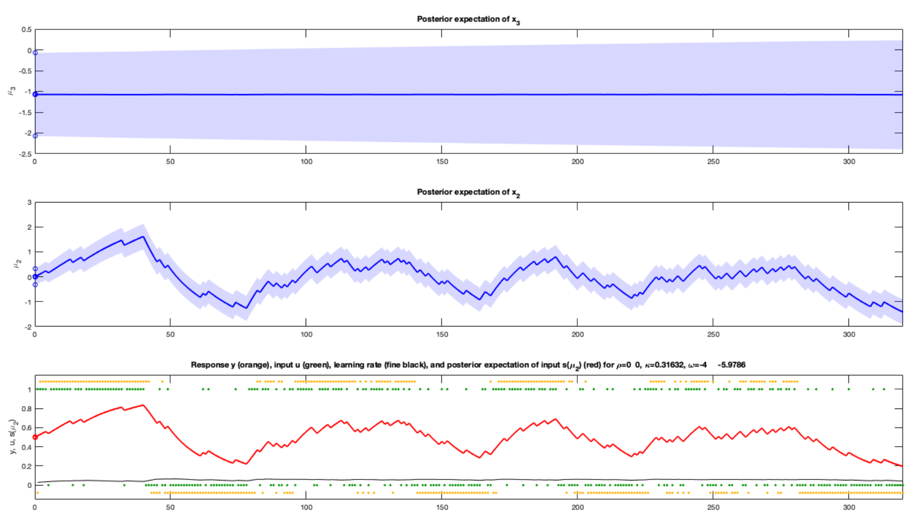
```
Ignored trials: none
Ignored trials: none
Irregular trials: none

Optimizing...

Calculating the log-model evidence (LME)...

Results:
```

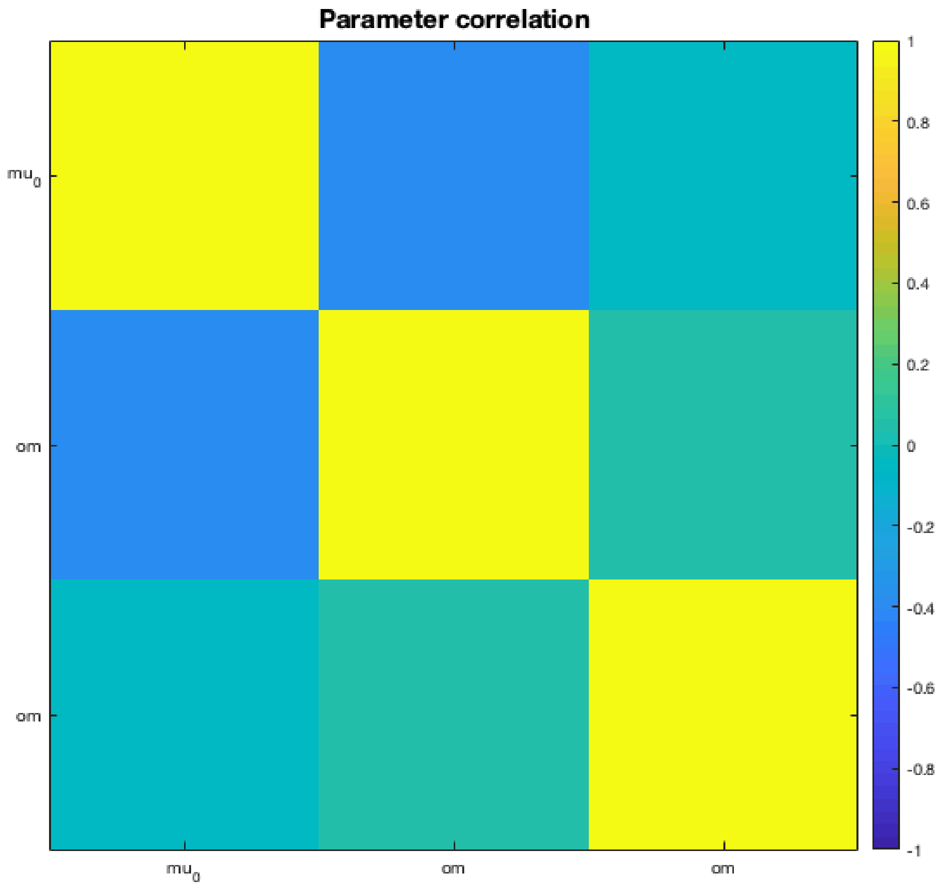
```
Parameter estimates for the perceptual model:
    mu_0: [NaN 0 -1.0744]
    sa_0: [NaN 0.1000 1]
     rho: [NaN 0 0]
      ka: [1 0.3163]
      om: [NaN -4 -5.9786]

Model quality:
    LME (more is better): -150.577
    AIC (less is better): 284.8375
    BIC (less is better): 296.1425

    AIC and BIC are approximations to -2*LME = 301.1539.


est1 =

  struct with fields:

        y: [320×1 double]
        u: [320×1 double]
      ign: []
      irr: [0×1 double]
    c_prc: [1×1 struct]
    c_obs: [1×1 struct]
    c_opt: [1×1 struct]
    optim: [1×1 struct]
    p_prc: [1×1 struct]
    p_obs: [1×1 struct]
     traj: [1×1 struct]

Ignored trials: none
Irregular trials: none

Optimizing...

Calculating the log-model evidence (LME)...

Results:

Parameter estimates for the perceptual model:
    mu_0: [NaN 0 -1.4751]
    sa_0: [NaN 0.1000 1]
     rho: [NaN 0 0]
      ka: [1 1]
      om: [NaN -0.1441 -5.9283]

Model quality:
    LME (more is better): -67.9986
    AIC (less is better): 124.8677
    BIC (less is better): 136.1726

    AIC and BIC are approximations to -2*LME = 135.9972.
```

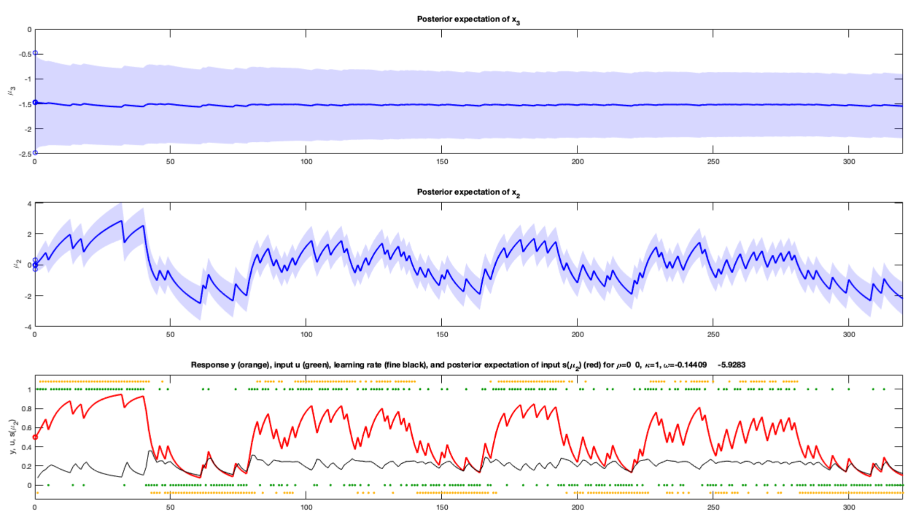
```
est2 =

  struct with fields:

        y: [320×1 double]
        u: [320×1 double]
      ign: []
      irr: [0×1 double]
    c_prc: [1×1 struct]
    c_obs: [1×1 struct]
    c_opt: [1×1 struct]
    optim: [1×1 struct]
    p_prc: [1×1 struct]
    p_obs: [1×1 struct]
     traj: [1×1 struct]
```




Parameter correlation
mu_0
ka
om
mu_0
ka
om
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1


Posterior expectation of x_3
μ_3
Posterior expectation of x_2
μ_2
Response y (orange), input u (green), learning rate (fine black), and posterior expectation of input s(μ_2) (red) for ρ=0 0, κ=0.31632, ω=-4 -5.9786
y, u, s(μ_2)
Trial number





Published with MATLAB® R2018b